DOCUMENT RESUME

ED 099 411 TM 004 069

AUTHOR Roudabush, Glenn E.; Green, Donald Ross
TITLE Aspects of a Methodology for Creating Criterion-Referenced Tests.
PUB DATE [Apr 72]
NOTE 50p.; Paper presented at the Annual Meeting of the National Council for Measurement in Education (Chicago, Illinois, April 1972)

EDRS PRICE MF-$0.75 HC-$3.15 PLUS POSTAGE
DESCRIPTORS Behavioral Objectives; *Criterion Referenced Tests; Diagnostic Tests; *Statistical Analysis; Student Evaluation; *Test Construction; *Test Reliability; Test Results
IDENTIFIERS Prescriptive Mathematics Inventory; Prescriptive Reading Inventory

ABSTRACT

In determining how reliable is reliable enough and how much error can be tolerated in criterion-referenced testing, the following relationships hold: (1) the more specific an objective is, the fewer the items required to reliably measure it; (2) the more specific the objectives are, the more objectives required to cover a given span of the curriculum; and (3) the more specific the objectives are, the more diagnostic the test is in terms of uncovering specific deficiencies in a student's achievement. There is also a four-way interdependency between comprehensiveness, measurement reliability, test length, and curriculum span. Data are presented and discussed with particular emphasis on the reliability problem and specificity-generality dimension of behaviorally stated objectives. Data were collected on the Prescriptive Mathematics Inventory and the Prescriptive Reading Inventory. It is concluded that more information is obtained about a student's specific strengths and weaknesses by measuring a large number of specific objectives with few items than by measuring a few general objectives with a large number of items. (Author/RC)

ED 099411

ASPECTS OF A METHODOLOGY FOR CREATING CRITERION-REFERENCED TESTS

by

Glenn E. Roudabush and Donald Ross Green

CTB/McGraw-Hill

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY.

"PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL HAS BEEN GRANTED BY
GLENN ROUDABUSH & DONALD ROSS GREEN
TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE NATIONAL INSTITUTE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER."

A paper presented at the

National Council for Measurement in Education meetings in Chicago

April 1972

TM 004 069

A criterion-referenced test is a test based on a set of specific learner objectives stated in terms of the behavioral changes to be expected in the examinee as a result of instruction toward these objectives. If you wish to make the objectives the basis for an objective test with general applicability that covers a reasonable span of the curriculum in a given content area, then a number of difficult problems arise: (1) how specific should the objectives be; (2) how important is a given objective to achievement in the subject area; (3) how can the objectives be made measurable with a paper and pencil test; (4) how many test items should be written to measure each objective; and (5) what is an acceptable standard of performance? With respect to the total set of objectives to be included in the test, comprehensive coverage of the span of curriculum desired and the structure or organization of the objectives will also be of concern. All these problems need to be solved, and in the end the final product must not be unreasonably long or time-consuming to administer. These problems are, of course, interrelated.

An item on a paper and pencil test, when completed by a student, is a sample of that student's behavior. Furthermore, the item is a sample selected from some universe of all possible behaviors which might have been selected to represent or measure some particular domain of behaviors. In a very real sense, the objective that the item is written to measure defines this universe of possible behaviors. The more limited the universe, the easier it is to select behaviors to represent it, and the more confidence we can have that the selected behaviors do, in fact, represent the domain.

The difficulty of establishing an acceptable standard of performance is directly related to the general problem of reliability and validity of measurement, which in turn concerns the extent to which the domain of behaviors specified by the objective is sampled. There are, then, two basic considerations. First, how many instances of the behavior constitutes a reliable measure of the domain; and second, of these, how many must a student perform correctly in order to be considered proficient in that class of behaviors? The generality-specificity dimension of behaviorally stated objectives is directly related to these considerations.

If the assessment device is an objective test, perhaps scored by computer, then the only information available is correct versus incorrect responses to individual items. If we assume some Platonic truth about the student with respect to his ability to perform the behaviors described by an objective and also assume a level of specificity of the objective such that (ideally) if a student can perform correctly one behavior from the domain, then he can perform them all, then for any given item on a test there are four possible outcomes: (1) the student cannot perform the behavior and does not get the item correct; (2) the student cannot perform the behavior but does get the item correct; (3) the student can perform the behavior but does not get the item correct; and (4) the student can perform the behavior and does get the item correct. If a large number of students were repeatedly tested with items sampling the domain of behaviors specified by such an objective, the proportion of correct responses of individual students should approach a stable bimodal distribution. The lower mode would give the mean proportion for students who cannot perform the behavior and the upper mode would give the mean

proportion for students who can perform the behavior.

The assumptions above are never true. We have ignored partial knowledge, and no objective is so specific that there is no variation in the difficulty of the behaviors it specifies. These two effects are confounded in any testing situation. There are no reliable measures. The questions then become how reliable is reliable enough and how much error of each kind are we willing to tolerate? The following relationships hold: (1) the more specific an objective is, the fewer the items required to reliably measure it; (2) the more specific the objectives are, the more objectives required to cover a given span of the curriculum; and (3) the more specific the objectives are, the more "diagnostic" the test is in terms of uncovering specific deficiencies in a student's achievement. There is also a four-way interdependency between comprehensiveness, reliability of measurement, test length, and curriculum span. In developing a criterion-referenced test for general use, all of these elements must be balanced and compromises made so that the resulting instrument is usable and useful.

All of the problems mentioned here cannot be discussed at length in this short paper. In what follows, data will be presented and discussed with particular emphasis on the reliability problem and the specificity-generality dimension of behaviorally stated objectives. The data were collected on the *Prescriptive Mathematics Inventory* (PMI) by John Gessel, published by CTB/McGraw-Hill, and the *Prescriptive Reading Inventory* (PRI), which is to be published later this year.

The author of the PMI, John Gessel, made an early decision to use a single item to measure an objective. In doing this, he was sacrificing

reliability of measurement in order to maintain a desired level of specificity of the objectives and a desired comprehensiveness and span of the curriculum. The decision was deliberate in order to make the information provided by the test as specific, and therefore diagnostic, as possible. We have collected and analyzed data on the PMI in order to determine the effect of using what a measurement theorist would call a "one-item test" to measure an objective. Logically, there are reasons to believe that the use of one item is not as serious a problem in the PMI as it appears on the surface. There are two ways in which an item can give erroneous information: the first is the "false positive," in which the student answers an item correctly, but has not mastered the underlying behavior; and the second is the "false negative," in which the student misses the item even though he has in fact mastered the underlying behavior.

In the PMI, the false positive response, which generally involves guessing, is virtually eliminated in that the format demands a constructed response which is then encoded in a grid for machine scoring. Figure 1 (Appendix A, p. A-1) shows an example item from the PMI and its associated response grid. The chance factor for most items is quite small, 1/100 or less in most cases. There are some items where the chance factor is 1/10 and a few where it is as large as 1/5, but even here the set of the student in taking this test is such that he is unlikely to mark the grid at random. Note that there are 1,000 ways to mark this grid. If the student marks the correct answer in the grid, then we can be reasonably sure that he knows the answer.

The false negative response is more difficult to handle. As long as the total test is seen simply as a large collection of one-item tests, then the value of the information provided will be suspect. In order to justify the use of one item per objective, we must look beyond the single item and the particular objective it purports to measure and consider the structure and organization of the test as a whole. The objectives measured by the test and the items written to measure them are generally hierarchically ordered from less to more complex, and the obtained item difficulties confirm this organization within major categories. In interpreting the results of testing, the user is urged to look at the pattern of right and wrong responses and to discount isolated wrong responses when they do not fit into an overall hierarchical pattern. Figure 2 (Appendix A, p. A-1) shows a section of the PMI Diagnostic Matrix, which is one of the reports returned to the student and his teacher. Notice the hierarchy implicit in the arrangement of the report and the pattern of responses of the hypothetical student represented in the illustration. There are several objectives (items) missed by the student that could probably be ignored by him and his teacher.

If you were constructing a norm-referenced test to measure, say, arithmetic computation, you would probably write a comprehensive rationale describing the kind of test you want, the kind of items you want, and perhaps you would provide a few sample items to give to your item writers. After the items were written and tryout data obtained, you would choose items whose item difficulty level hovers around .5, because they discriminate best, and which are internally consistent. The assumption is that there

is an underlying continuum called "arithmetic computation achievement" and you wish to place students on this continuum as accurately as possible. Furthermore, you would want the scores obtained from your test, when administered to the target population, to be distributed approximately as the normal distribution. Such test scores have valuable psychometric properties: they correlate well with other similar scores, provide meaningful derived scores, and so on.

In constructing a criterion-referenced test, the items written to measure a single, specific objective in that test do not share the desirable properties described for a norm-referenced test. In a criterion-referenced test it is not necessarily desirable for the item difficulties to be near .5, nor is internal consistency especially important. In particular, a normal distribution of the objective scores is distinctly not wanted. There is no assumption of an underlying continuum of achievement on a particular objective, but rather there is the simple assumption that each student either has mastered the objective or that he has not mastered the objective. The desired distribution of scores on the one objective when administered to the target population is bimodal or U-shaped as shown in the hypothetical distribution in Figure 3 (Appendix A, p. A-2). Desirable hypothetical distributions for an easy test and a hard test are shown in Figures 4 and 5 (Appendix A, p. A-2), respectively, and a hypothetical distribution where learning has apparently occurred during the testing session is shown in Figure 6 (Appendix A, p. A-2). Examples of each of these distributions based on real data will be described shortly.

The PMI reliability study from which these examples were derived was carried out last year in two school districts, one in northern, the other in southern California. In these schools, we administered Level B of the PMI to seventh graders along with a set of criterion tests, and we administered Level C of the PMI to eighth graders along with another set of criterion tests. To construct the criterion tests, we first selected 40 items from the PMI to represent the range of material covered by all levels of the test. We then wrote alternate forms of these items to form either 5-item or 20-item criterion tests to measure each of the 40 objectives represented by the selected PMI items. This selection of items was done from the tryout version of the test and the final version omitted some of these 40 items. We also discarded results for two additional tests for which less than 10 students correctly answered the corresponding PMI item. The results reported here, then, represent 32 rather than 40 objectives from Levels B and C of the PMI. The list of objectives measured by these criterion tests is shown in Table 1 (Appendix B, p. B-1). Summary statistics for these criterion tests are given in Table 2 (Appendix B, p. B-4). Also included in the table is the point biserial correlation between the criterion test score and its corresponding PMI item. Notice that this correlation tends to fall somewhat below the lowest point biserial for any item in the criterion test itself, though in a few cases it falls within the range. The week separation between the administration of the PMI item and the criterion tests may account for this difference.

Now look at Figure 7 (Appendix A, p. A-3). This figure shows the obtained distribution for criterion test B-2, a 20-item test. The solid

line represents the total distribution, the dashed line represents the distribution of the criterion test scores for those students who correctly answered the corresponding PMI item, and the hatched line represents the distribution for those students who incorrectly answered the corresponding PMI item. This distribution closely resembles the hypothetical distribution for an appropriate test shown in Figure 3 (Appendix A, p. A-2). Another example of the same form of distribution, this time for a 5-item criterion test, is shown in Figure 8 (Appendix A, p. A-4). Figures 9 (Appendix A, p. A-5) and 10 (Appendix A, p. A-6) show distributions for 20-item and 5-item criterion tests, respectively, following the pattern for the hypothetical easy test shown in Figure 4 (Appendix A, p. A-2). Figures 11 (Appendix A, p. A-7) and 12 (Appendix A, p. A-8) show the 20-item and 5-item criterion test distribution following the hypothetical distribution for a difficult test shown in Figure 5 (Appendix A, p. A-2). And, finally, Figure 13 (Appendix A, p. A-9) and 14 (Appendix A, p. A-10) show 20-item and 5-item criterion test distributions conforming to the hypothetical distribution shown in Figure 6 (Appendix A, p. A-2) where learning has occurred for some students. All of the criterion test distributions can be classified into one or another of these hypothetical forms except for one, B-4, which could not be classified. The classifications of all the tests are shown in Table 2 (Appendix B, p. B-4), along with the statistics mentioned earlier.

Figure 14 (Appendix A, p. A-10) is especially interesting, since it shows a considerable amount of what we have called learning by students who missed the corresponding PMI item for this objective. Figure 15 (Appendix A, p. A-11) gives an example of the kind of item used in this

test and its response grid. Note that the three digits required for the answer are found by working from right to left in the problem and that the response grid indicates that they are to be recorded in the right to left order. In the PMI itself, about 8% of the students gridded their response in reverse order. Approximately the same percentage of students gridded in reverse order for the criterion test items. Apparently the learning cannot be attributed to increased facility in gridding the responses. The item difficulty of the corresponding PMI item, however, is substantially lower than the difficulties of any of the criterion test items, about .46 as opposed to .63 to .66. The only obvious difference in the original item was that the size of the blanks (square, triangle, etc.) was larger in the PMI item. Our tentative conclusion is that, in the process of trying 212 PMI items one week and 196 similar criterion test items the next week, a substantial number of students, in fact, learned how to work the problem.

In a second phase of the PMI reliability study, we postulated that a better estimate of the mastery of specific objectives could be obtained by using more of the information contained in the total test. Since the objectives measured are not independent of each other, we designated subscores for the PMI, and using a stepwise regression procedure we used these subscores to predict each of the criterion test scores in turn. Table 3 (Appendix B, p. B-5 ff.) gives a short description of each of the predictor subscores defined for this study and the number of items contained in each of them. The scores designated "B" were used to predict the Level B criterion tests and those designated "C" were used for the

Level C criterion tests. The stepwise procedure was terminated when the next predictor to be added would increment the multiple correlation by less than .003. In every case, using all of the predictors would not have increased the multiple correlation by more than about .01, and in most cases the shrunken multiple correlation would have decreased.

A summary of this analysis is shown in Table 4 (Appendix B, p. B-9), giving the number of predictors selected, the order of selection, the multiple correlation, and shrunken multiple correlation. For comparison, the point biserial correlations of the corresponding PMI items with the criterion tests are also shown. In every case except one, the multiple correlation is substantially higher than the point biserial with the corresponding PMI item. Criterion test B-9, having to do with the intersection of sets, seems not to be predictable from the PMI subscores as defined here. Where the point biserial correlation is small, the difference between it and the multiple correlation tends to be large, but where the point biserial is large, the difference tends to be small. This may reflect a ceiling effect and may also be related to how well the particular criterion test measures its objective. No cross validation was carried out and, of course, the validity coefficients that would be obtained would be smaller than the multiple correlations shown.

In most cases the single item seems to give a good indication of mastery or non-mastery of its objective. Where the point biserial correlations are low, the cause can generally be traced to either a somewhat poor set of items in the criterion test, or to those tests in which learning effects seem to be present. The substantial added effort and cost of obtaining regression estimates for all of the objectives in the PMI would probably not

substantially improve the reliability of the information provided by the instrument.

The PRI reliability study was designed and carried out in a somewhat different way than was the PMI study. Separate criterion tests were not written for the PRI. Instead, as part of the item tryout study, ten or more items were written to measure certain objectives. These objectives were designated as "reliability study objectives" or PRI criterion tests. In some cases, similar objectives were combined into one in order to get the needed number of items for this study. The list of PRI criterion tests is shown in Table 5 (Appendix B, p. B-10 ff.). They are designated A, B, C, or D in accordance with the level of the test in which they appear. This study was further complicated by the large number of items to be tried out. Each level of the test was split into several booklets: Level A into eight booklets, and Levels B, C, and D into four. Each student in the study took only one of the 20 separate booklets.

In most cases, the distributions for the objective scores from the PRI did not show the desired U-shaped curve that we found with the PMI criterion tests. Of approximately 400 obtained distributions, only 19 clearly showed the U-shaped form. An additional 16 showed some indication of the U-shaped distribution, but most of the distributions either approximated a normal distribution or were more or less severely skewed to the left or right. In an effort to measure as purely as possible the defined behaviors appearing in the test, we kept stringent control over the vocabulary level. As a result, the test was generally easy for the students taking it, and for this reason more than half of the distributions were classified as strongly skewed to the right.

Figure 16 (Appendix A, p. A-11) shows our best example of a U-shaped distribution from the PRI. This objective had only five items in the test booklet in which it appeared, so it was not part of the reliability study. It does look remarkably like some of the distributions found for the PMI criterion tests, for example, Figure 8 (Appendix A, p. A-4). Note that it is displaced upward, probably reflecting the fact that these are multiple choice items and the guessing factor is in evidence.

The intent of the PRI reliability study was to discover the minimum number of multiple choice items that could reasonably be used to measure an objective. The procedure used was to select two items at random from the pool of ten or more items for that criterion test objective, obtain the product moment correlation between the score on those two items and the score on the remaining items, and obtain the joint frequency distribution between the two scores. A third item was then chosen at random and the correlation and joint frequency distribution re-computed; then a fourth item was added, and so on until either a correlation of .80 was obtained or until half of the available items were chosen.

The results of this analysis show that in most cases, three or four items are sufficient to give a reasonably good estimate of mastery or non-mastery of the objective. Figure 17 (Appendix A, p. A-12) shows the best example where three items are used from a pool of thirteen and where the criterion of mastery is specified as two of the three items correct. This distribution follows the pattern of an easy test given in the hypothetical example shown in Figure 4 (Appendix A, p. A-2). Those students meeting the criterion of two out of three items account for the high scores, and those students not meeting the criterion account

for the low scores. The higher scoring students who did not meet the criterion can be accounted for by the multiple choice format of the test. For example, 14 students would be expected to score four on the test by chance, while 18 actually did. On the other hand, less than the expected number of students not meeting the criterion scored three on the test.

Figure 18 (Appendix A, p. A-13) shows an example of one of the poorer distributions. Here three items are again used and the criterion specified is two of the three correct. The total number of items in the pool was 12, so the distribution is for a nine-item test. Even here the high scores are again accounted for by those students meeting the criterion and the low scores by the students not meeting the criterion, though the pattern is not quite as clear.

Table 6 (Appendix B, p. B-15 ff.) gives a summary of the PRI criterion test statistics and the results of correlating the subsets of items chosen at random with the remaining items for each test. The reading tests are not as homogeneous as the math tests, reflecting the greater generality of the reading objectives. In many cases, however, the correlations for three or four items with the remaining items for a criterion test objective are respectable and it appears quite practical to use three or four items to measure these objectives.

As in the PMI reliability study, a second phase of the PRI study was carried out in which a stepwise regression procedure was applied to PRI subscores to predict the criterion test scores. This analysis is not as informative as the PMI analysis because of the separation of the test into many booklets. As a result, there are few predictors available for

any one criterion test. The set of predictors arranged by the criterion tests to which they apply are described in Table 7 (Appendix B, p. B-18 ff.). A summary of the regression analysis results are shown in Table 8 (Appendix B, p. B-21) giving the number of predictors selected, the order of selection, the multiple correlation, and the shrunken multiple correlation. For comparison, the 4-item test correlations with the remaining items in the test are also shown. The multiple correlations are somewhat higher than the 4-item correlations, but on cross validation the validity coefficients would again be smaller. Little advantage is shown here for the regression approach. The PRI reliability study will be replicated this fall as part of a larger study, so the results reported here can be taken as tentative.

It is our conclusion that more information is obtained about a student's specific strengths and weaknesses in math or reading by measuring a large number of specific objectives with few items than by measuring a few general objectives with a large number of items, even though the statistical reliability of such scores is lower than one would want. In interpreting the results obtained from the PRI, as with the PMI, we will urge the student and his teacher to look at the overall pattern of scores displayed on a diagnostic map of reading objectives. Practically, although the individual indications of mastery or non-mastery of the objectives are fallible, few wrong decisions will be made about any student when the total pattern of responses is taken into account.

FIGURES AND TABLES TO ACCOMPANY THE PAPER ENTITLED "ASPECTS OF A METHODOLOGY FOR CREATING CRITERION-REFERENCED TESTS"

by

Glenn E. Roudabush and Donald Ross Green

CTB/McGraw-Hill

A paper presented at the annual meeting of the National Council for Measurement in Education, Chicago, April 1972

BEST COPY AVAILABLE

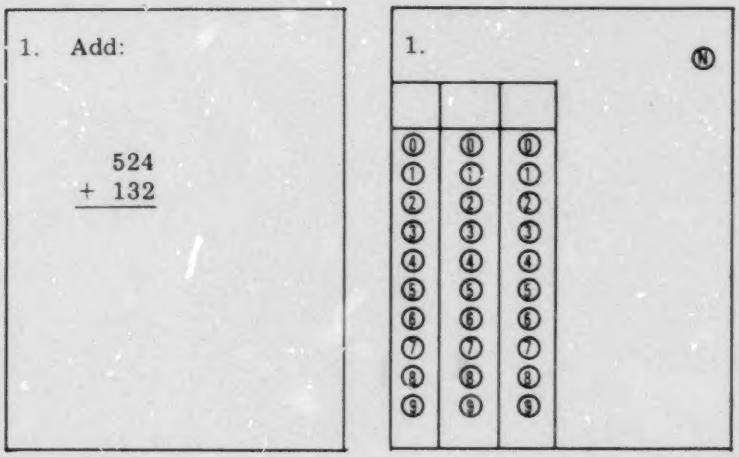

Figure 1. Example of a PMI Item and Its Response Grid.

PRESCRIPTIVE MATHEMATICS INVENTORY
DIAGNOSTIC MATRIX
LEVEL B

pi MATHEMATICS

OPERATIONS AND THEIR PROPERTIES		WHOLE NUMBERS				POSITIVE FRACTIONS				DECIMAL FRACTIONS			
		+	−	×	÷	+	−	×	÷	+	−	×	÷
NUMBER LINE		+	+	+		+	+	−	−				
ADDITION NO REGROUPING													
	PICTORIAL FRACTIONS					+	EQUIVALENCE			+			
	EQUIVALENCE					+	10THS			+			
	LOWEST (SIMPLEST) TERMS					+	100THS			+			
	LIKE					+	1000THS			+			
	UNLIKE					−							
	MIXED					+							
ADDITION REGROUPING	4 DIGIT	+	LIKE			+	10THS			−			
	5 DIGIT	+	UNLIKE			−	100THS			−			
	COLUMN	+	MIXED			−	1000THS			−			
SUBTRACTION NO REGROUPING						LIKE	+	10THS		+			
						UNLIKE	+	100THS		+			
						MIXED	+						
SUBTRACTION REGROUPING	2 DIGIT	+				LIKE	−	10THS		+			
	3 DIGIT	+				UNLIKE MIXED	−	100THS		−			
	4 DIGIT	+						1000THS		−			
	5 DIGIT	+											

CTB Published by CTB/McGraw-Hill, Del Monte Research Park, Monterey, California 93940. Copyright © 1971 by McGraw-Hill, Inc. All Rights Reserved. Printed in the U.S.A.

Figure 2. Section of the PMI Individual Diagnostic Matrix for Level B Showing Test Results for a Hypothetical Student.

HYPOTHETICAL DISTRIBUTION OF A CRITERION TEST FOR

1. A Total Sample, ———
2. A Portion of the Sample Correctly Answering the Corresponding Item of Interest, and - - - - - - -
3. A Portion of the Sample Incorrectly Answering the Corresponding Item. +++++++

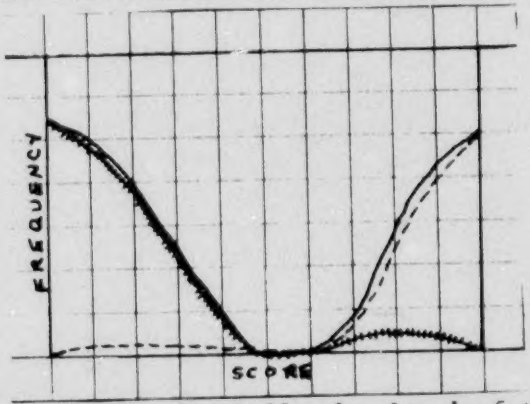

Figure 3. The Difficulty Level of the Test Is Assumed Appropriate for the Sample Tested. (Type A Distribution)

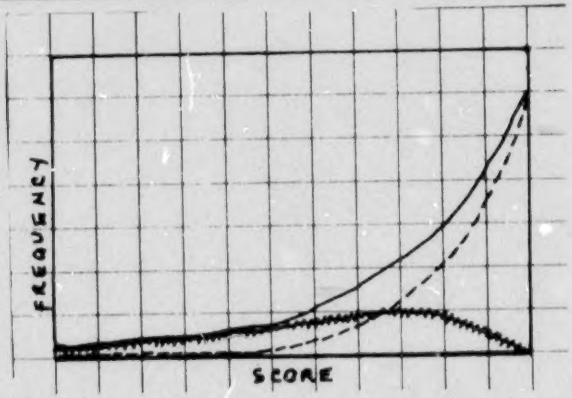

Figure 4. The Test Is Assumed Easy for the Sample Tested. (Type B Distribution)

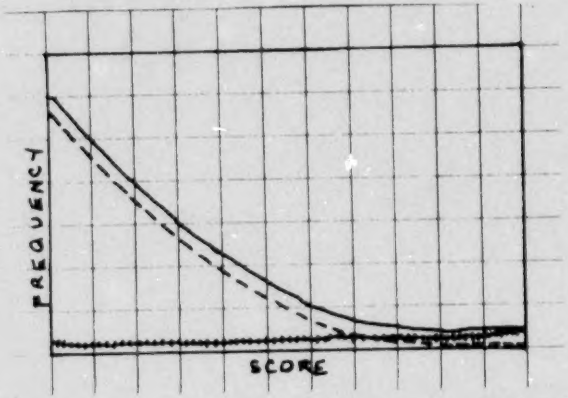

Figure 5. The Test Is Assumed Difficult for the Sample Tested. (Type C Distribution)

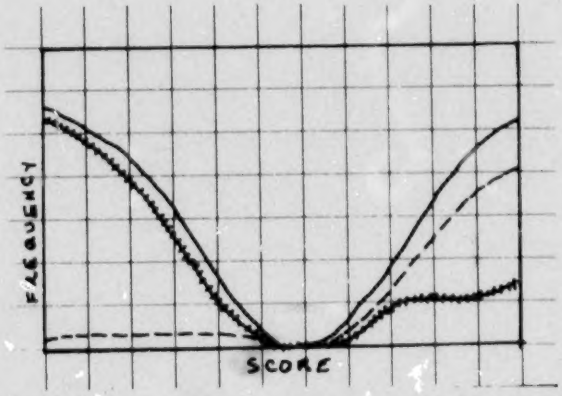

Figure 6. The Test Is Assumed Appropriate in Difficulty Level and Some Learning Is Assumed to Have Taken Place. (Type D Distribution)

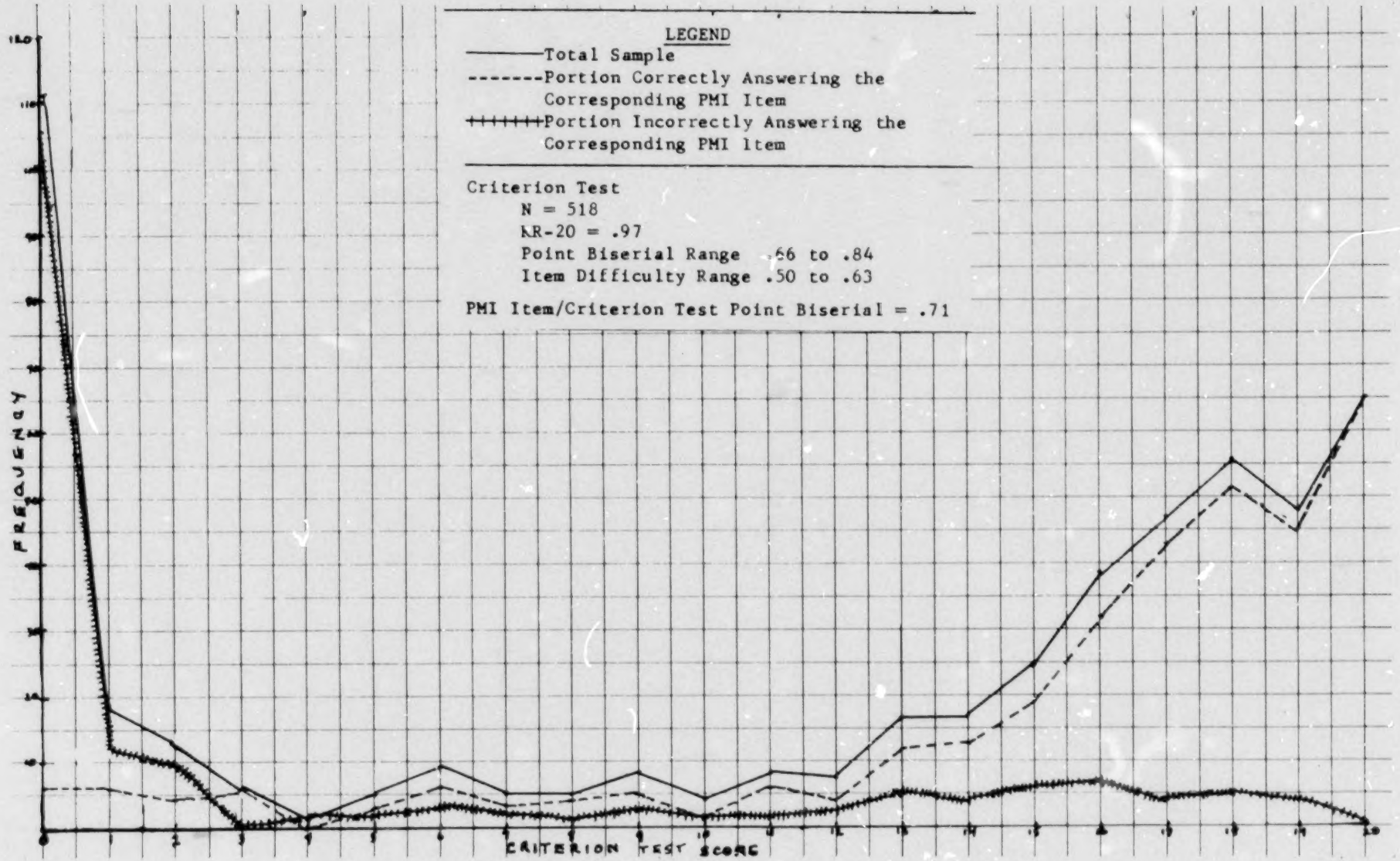
LEGEND
Total Sample
Portion Correctly Answering the Corresponding PMI Item
Portion Incorrectly Answering the Corresponding PMI Item
Criterion Test
N = 518
KR-20 = .97
Point Biserial Range .66 to .84
Item Difficulty Range .50 to .63
PMI Item/Criterion Test Point Biserial = .71
FREQUENCY
CRITERION TEST SCORE

Figure 7. Distribution of Scores on PMI Criterion Test B-2: ADDITION OF THREE POSITIVE FRACTIONS for (1) The Total Sample, (2) That Portion of the Sample Correctly Answering the Corresponding PMI Test Item, and (3) That Portion of the Sample Incorrectly Answering the Corresponding PMI Item.

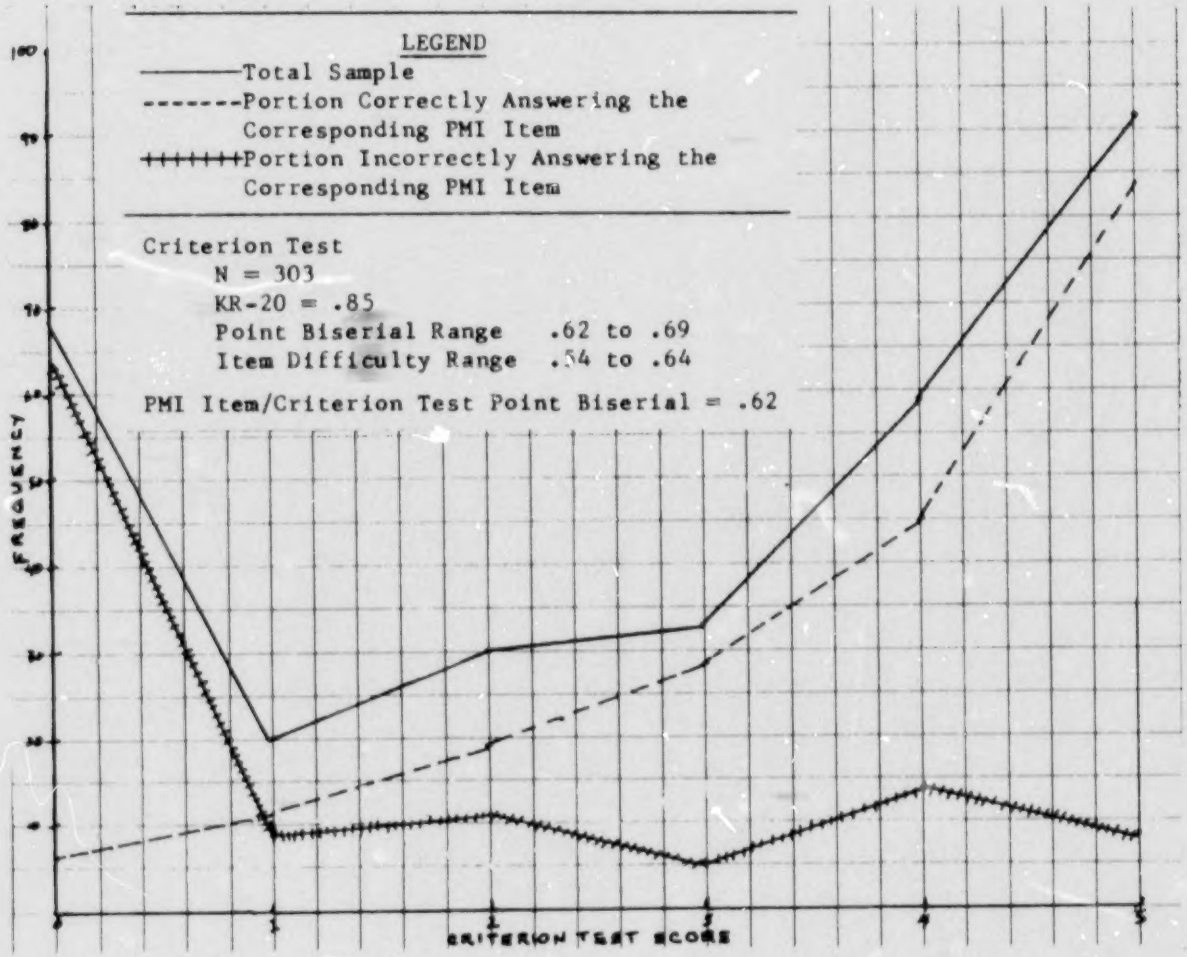
LEGEND
———Total Sample
-------Portion Correctly Answering the Corresponding PMI Item
+++++++Portion Incorrectly Answering the Corresponding PMI Item
Criterion Test
N = 303
KR-20 = .85
Point Biserial Range .62 to .69
Item Difficulty Range .54 to .64
PMI Item/Criterion Test Point Biserial = .62
FREQUENCY
CRITERION TEST SCORE

Figure 8. Distribution of Scores on PMI Criterion Test C-1: SUBTRACTION OF MIXED NUMBERS for (1) The Total Sample, (2) That Portion of the Sample Correctly Answering the Corresponding PMI Test Item, and (3) That Portion of the Sample Incorrectly Answering the Corresponding PMI Item.

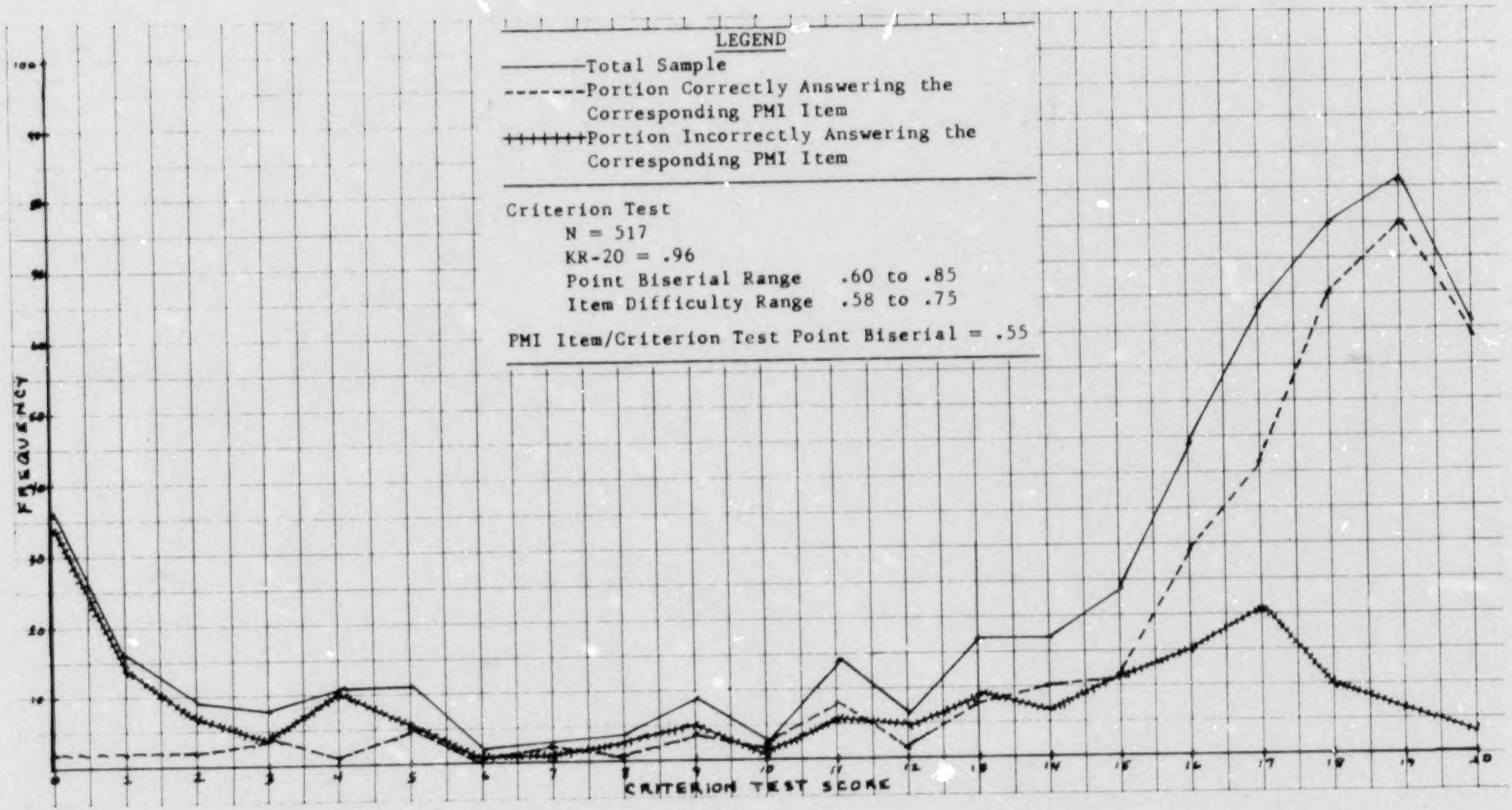
LEGEND
Total Sample
Portion Correctly Answering the Corresponding PMI Item
Portion Incorrectly Answering the Corresponding PMI Item
Criterion Test
N = 517
KR-20 = .96
Point Biserial Range .60 to .85
Item Difficulty Range .58 to .75
PMI Item/Criterion Test Point Biserial = .55
FREQUENCY
CRITERION TEST SCORE

Figure 9. Distribution of Scores on PMI Criterion Test B-5: DIVISION OF WHOLE NUMBERS WITH REMAINDER for (1) The Total Sample, (2) That Portion of the Sample Correctly Answering the Corresponding PMI Test Item, and (3) That Portion of the Sample Incorrectly Answering the Corresponding PMI Test Item.

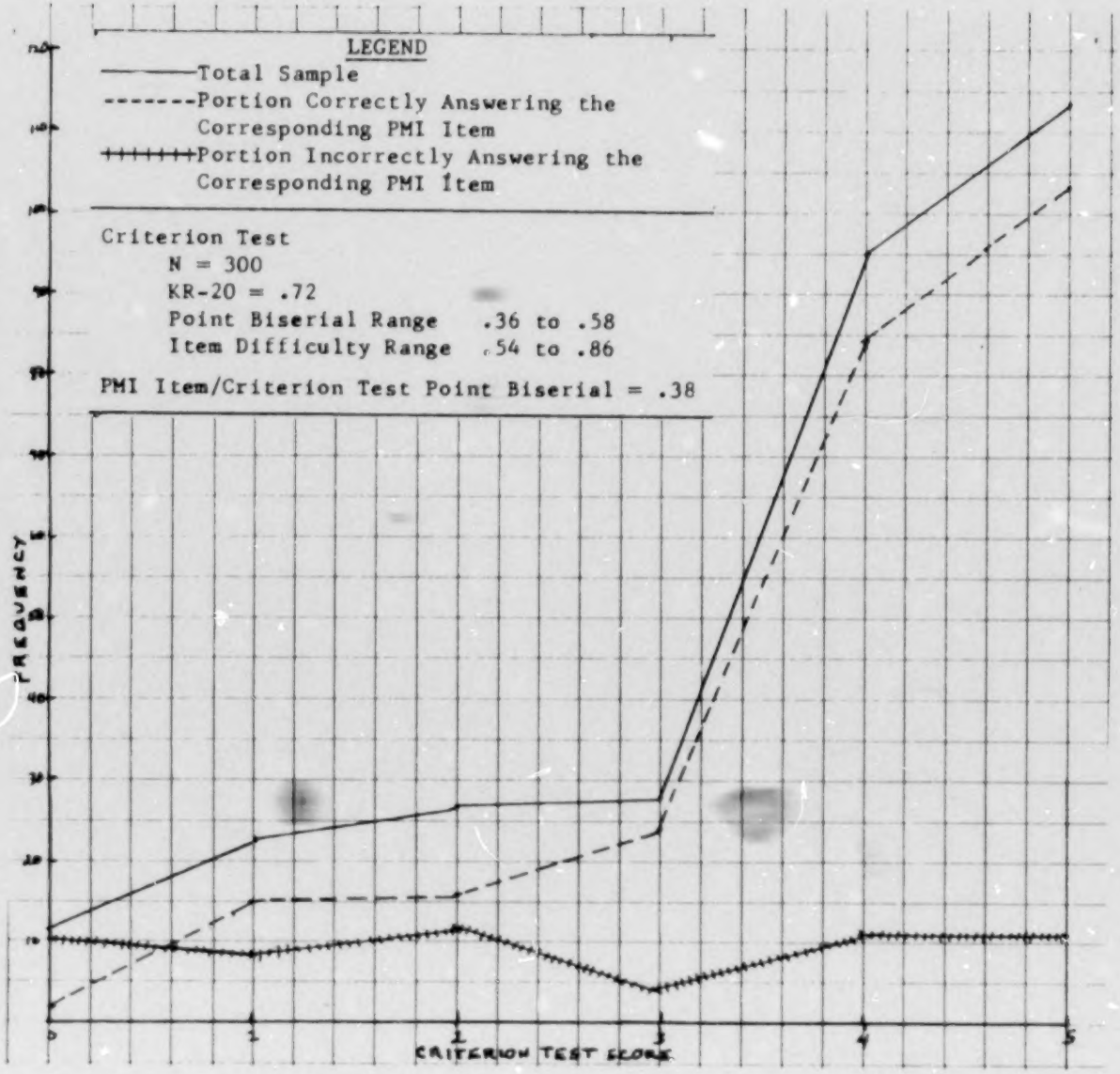

Figure 10. Distribution of Scores on PMI Criterion Test C-18: MAP READING (MODELS) for (1) The Total Sample, (2) That Portion of the Sample Correctly Answering the Corresponding PMI Test Item, and (3) That Portion of the Sample Incorrectly Answering the Corresponding PMI Item.

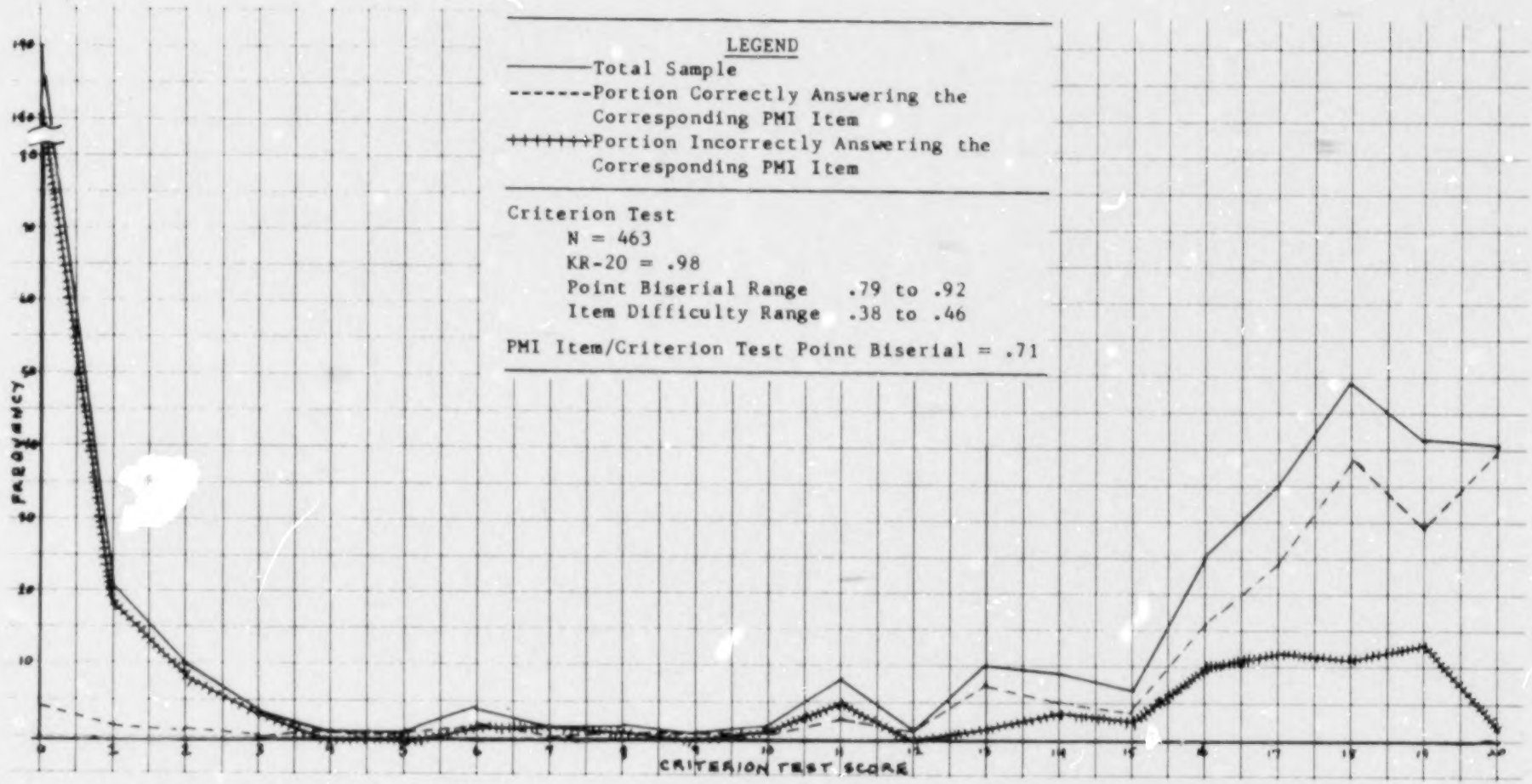
LEGEND
Total Sample
Portion Correctly Answering the Corresponding PMI Item
Portion Incorrectly Answering the Corresponding PMI Item
Criterion Test
N = 463
KR-20 = .98
Point Biserial Range .79 to .92
Item Difficulty Range .38 to .46
PMI Item/Criterion Test Point Biserial = .71
FREQUENCY
CRITERION TEST SCORE

Figure 11. Distribution of Scores on PMI Criterion Test B-6: MULTIPLICATION OF MIXED NUMBERS for (1) The Total Sample, (2) That Portion of the Sample Correctly Answering the Corresponding PMI Test Item, and (3) That Portion of the Sample Incorrectly Answering the Corresponding PMI Test Item.

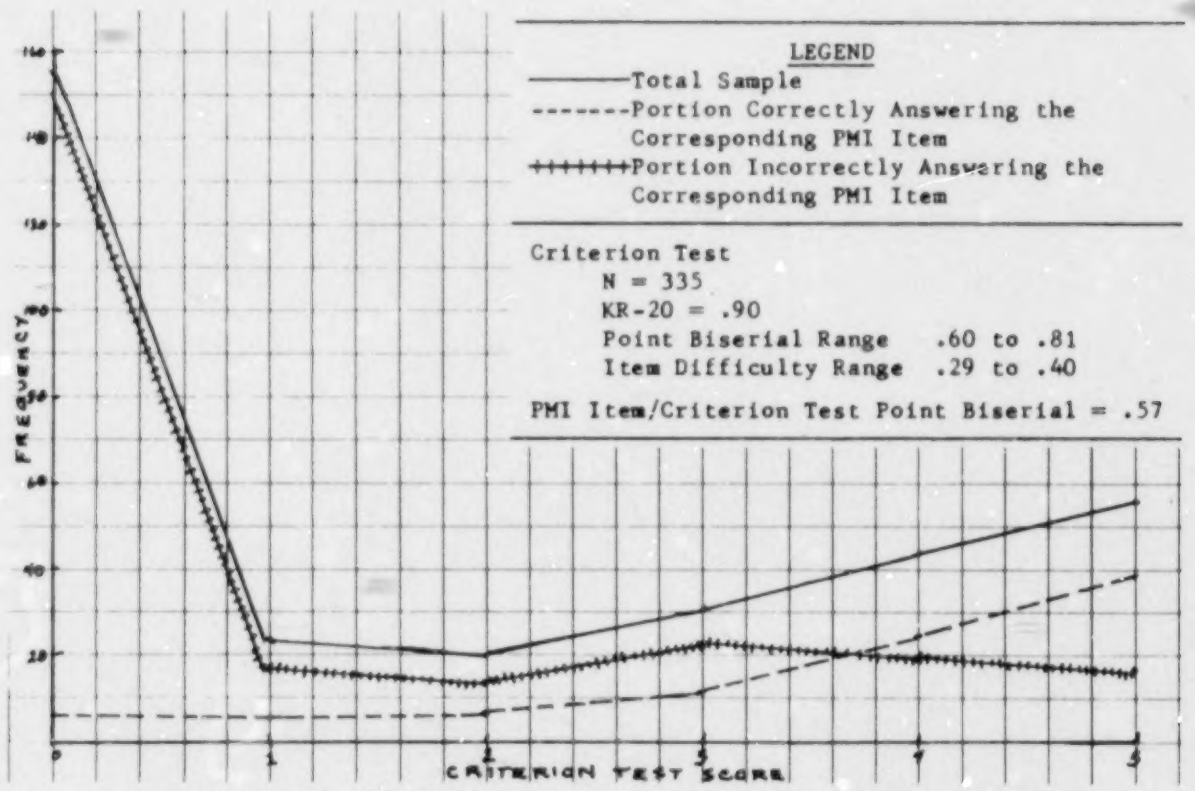
LEGEND
Total Sample
Portion Correctly Answering the Corresponding PMI Item
Portion Incorrectly Answering the Corresponding PMI Item
Criterion Test
N = 335
KR-20 = .90
Point Biserial Range .60 to .81
Item Difficulty Range .29 to .40
PMI Item/Criterion Test Point Biserial = .57
FREQUENCY
CRITERION TEST SCORE

Figure 12. Distribution of Scores on PMI Criterion Test B-16: DIVISION OF DECIMAL NUMBERS for (1) The Total Sample, (2) That Portion of the Sample Correctly Answering the Corresponding PMI Test Item, and (3) That Portion of the Sample Incorrectly Answering the Corresponding PMI Item.

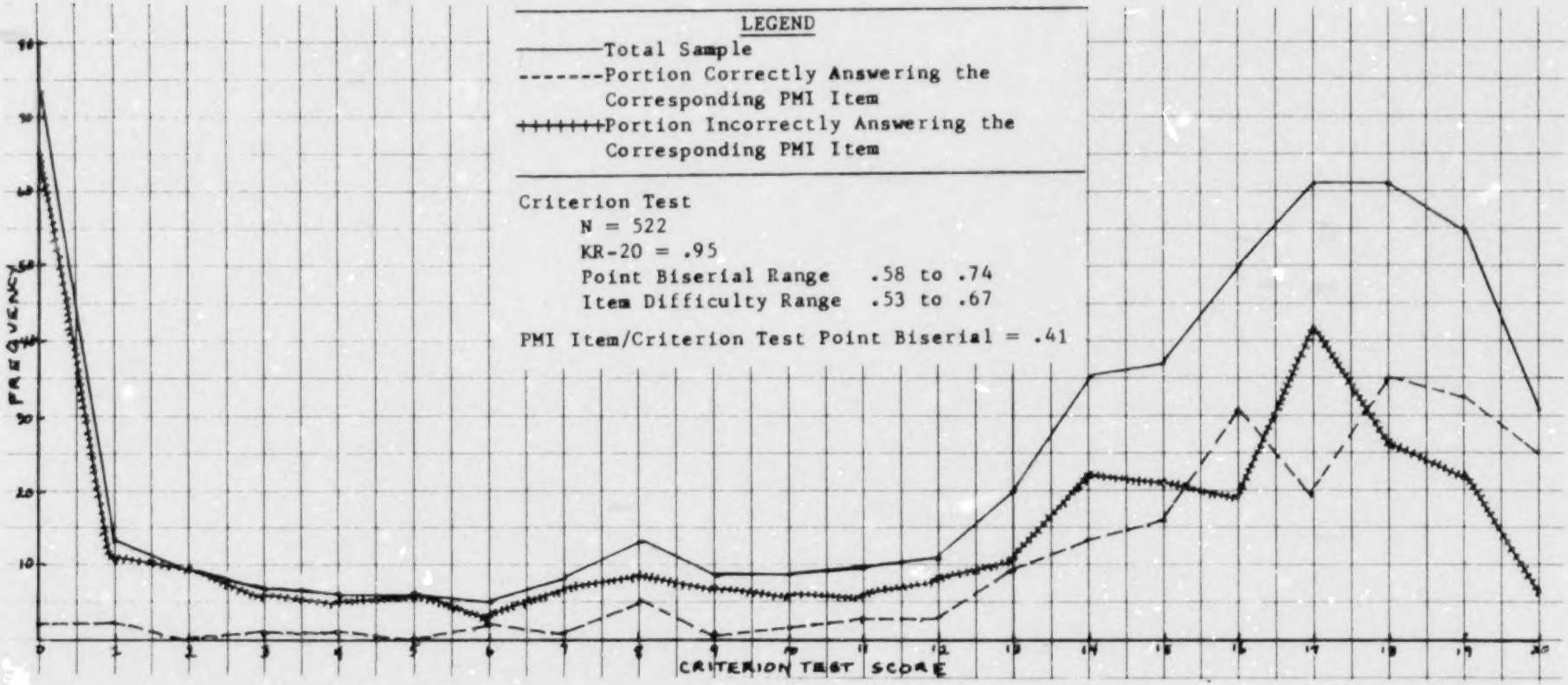
LEGEND
———Total Sample
-------Portion Correctly Answering the Corresponding PMI Item
+++++++Portion Incorrectly Answering the Corresponding PMI Item
Criterion Test
N = 522
KR-20 = .95
Point Biserial Range .58 to .74
Item Difficulty Range .53 to .67
PMI Item/Criterion Test Point Biserial = .41
% FREQUENCY
CRITERION TEST SCORE

Figure 13. Distribution of Scores on PMI Criterion Test B-8: ADDITION OF TIME for (1) The Total Sample, (2) That Portion of the Sample Correctly Answering the Corresponding PMI Test Item, and (3) That Portion of the Sample Incorrectly Answering the Corresponding PMI Test Item.

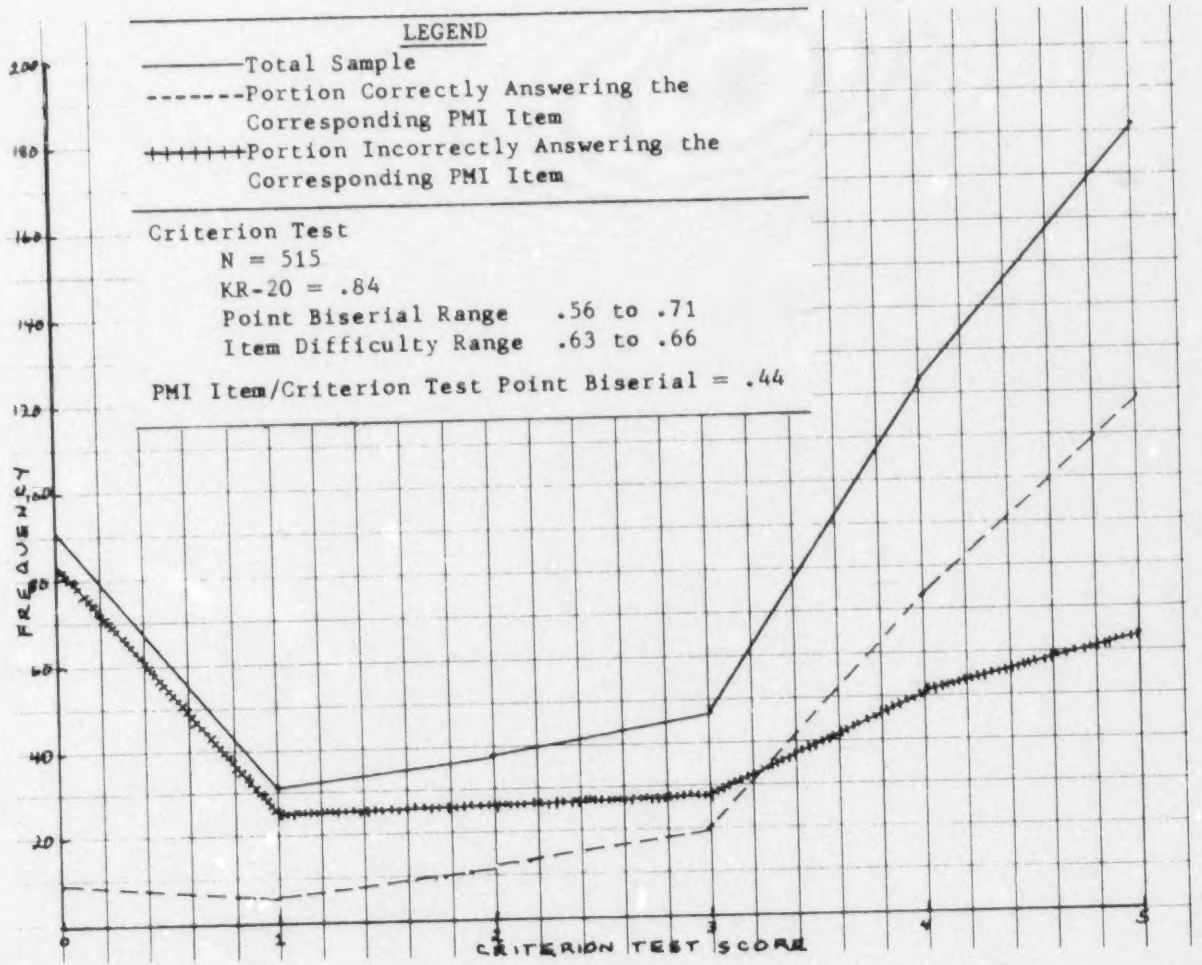
LEGEND
Total Sample
Portion Correctly Answering the Corresponding PMI Item
Portion Incorrectly Answering the Corresponding PMI Item
Criterion Test
N = 515
KR-20 = .84
Point Biserial Range .56 to .71
Item Difficulty Range .63 to .66
PMI Item/Criterion Test Point Biserial = .44
FREQUENCY
CRITERION TEST SCORE

Figure 14. Distribution of Scores on PMI Criterion Test B-19: MISSING DIGITS for (1) The Total Sample, (2) That Portion of the Sample Correctly Answering the Corresponding PMI Test Item, and (3) That Portion of the Sample Incorrectly Answering the Corresponding PMI Test Item.

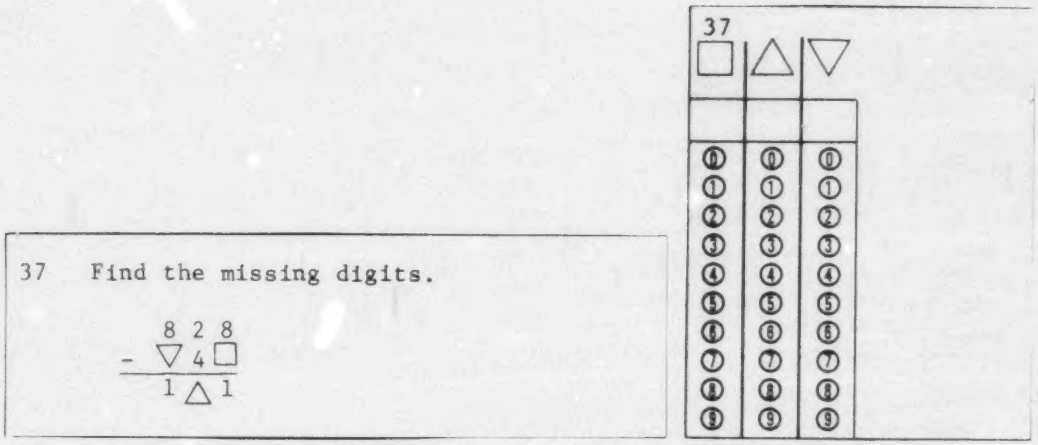

Figure 15. Example of the Kind of Item Used in Criterion Test B-19.

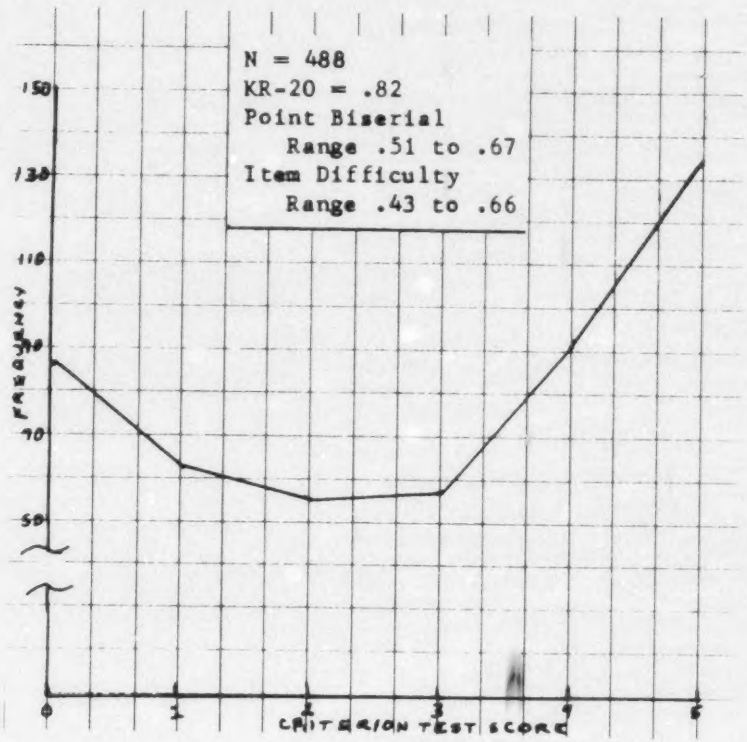

Figure 16. Distribution of Scores on PRI Objective 12 Level B: PHONETIC PARTS.

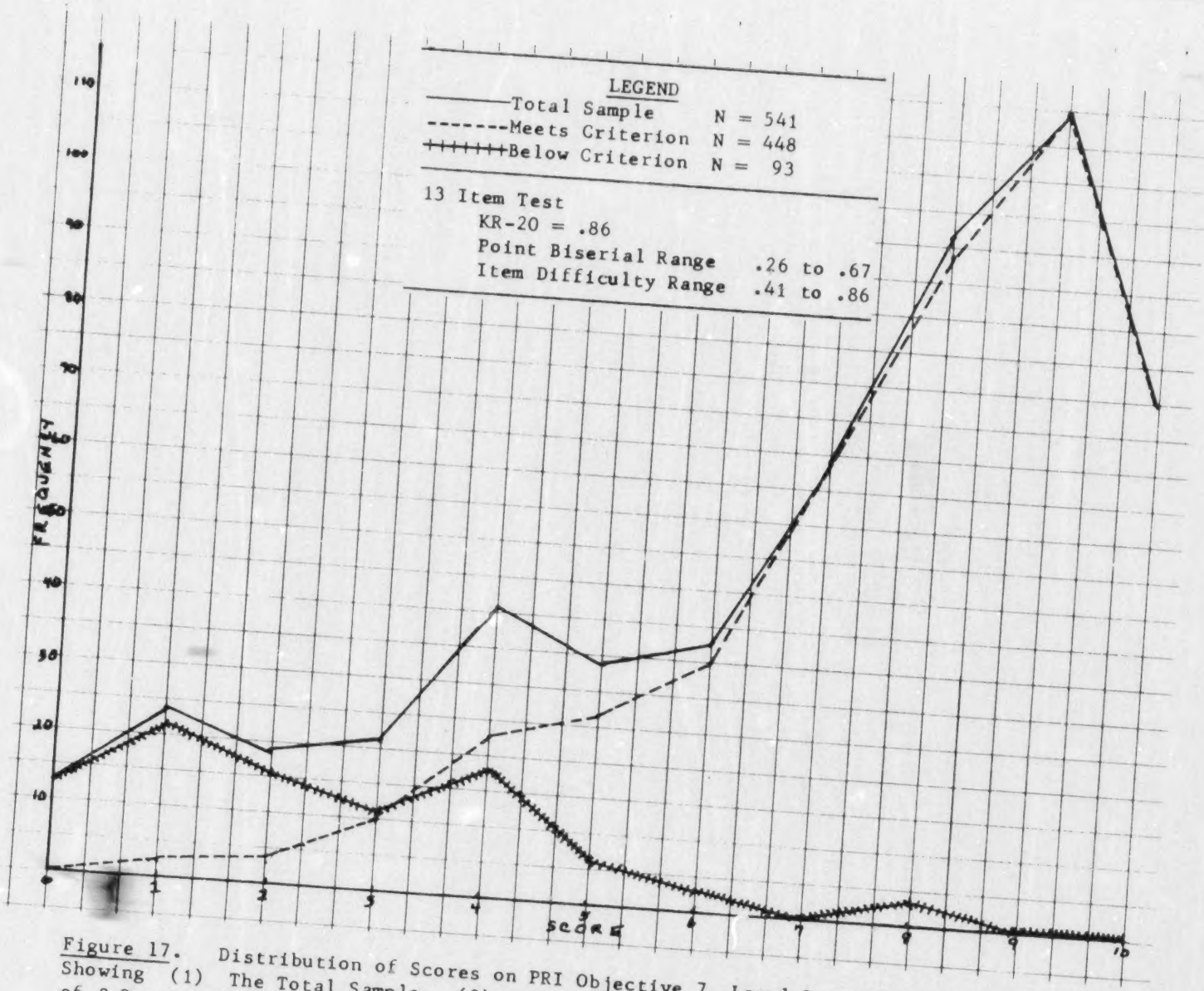

Figure 17. Distribution of Scores on PRI Objective 7, Level D: COMPOUNDS Showing (1) The Total Sample, (2) That Part of the Sample Not Meeting the Criterion of 2 Out of 3 Items Correct, and (3) That Part of the Sample Meeting the Criterion. The Correlation Between the 3 Randomly Selected Items and the Remaining 10 Items is .75.

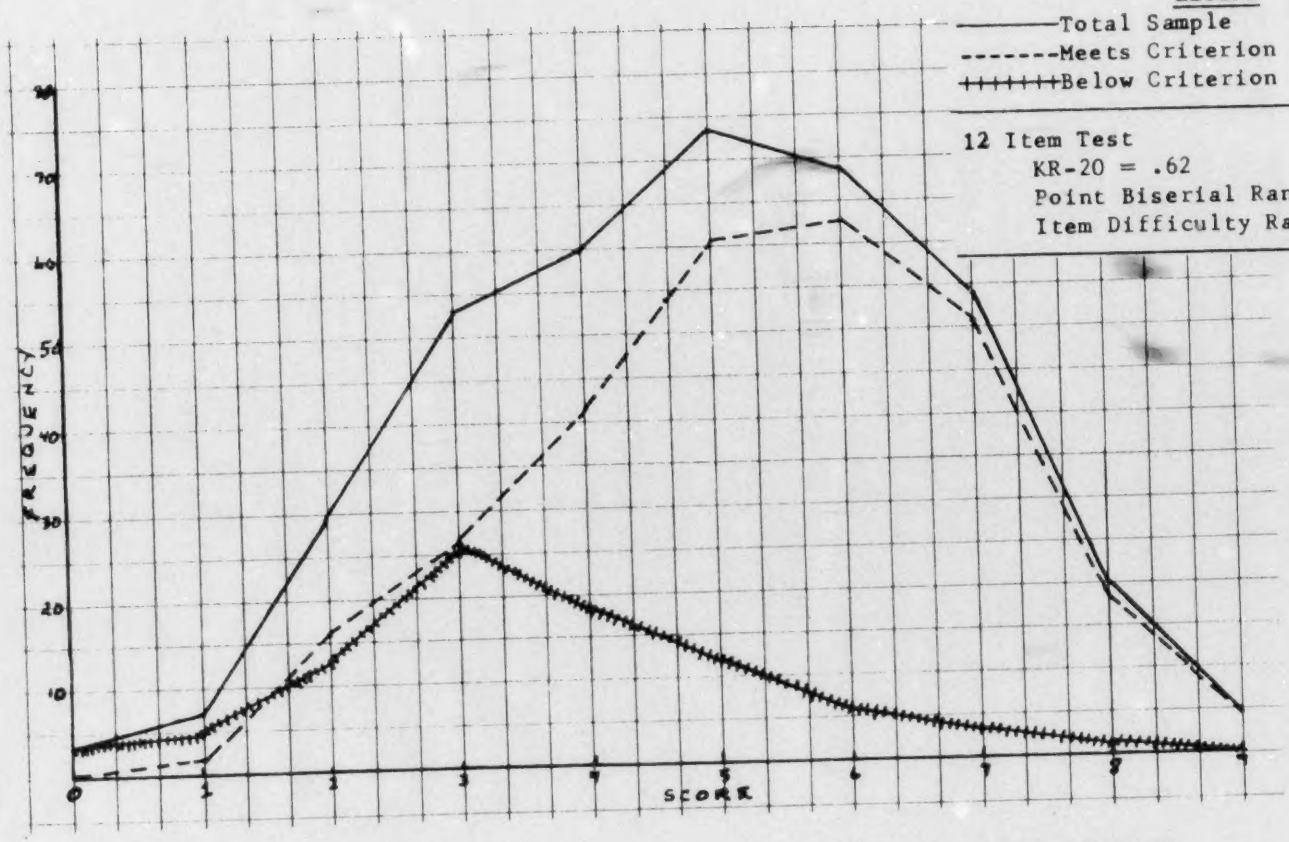

Figure 18. Distribution of Scores on PRI Objective 6, Level B: EVENT SEQUENCE Showing (1) The Total Sample, (2) That Part of the Sample Not Meeting the Criterion of 2 Out of 3 Items Correct, and (3) That Part of the Sample Meeting the Criterion. The Correlation Between the 3 Randomly Selected Items and the Remaining 9 Items is .46.

TABLE 1

OBJECTIVES MEASURED BY
THE 32 CRITERION TESTS CONSTRUCTED FOR THE
PMI RELIABILITY STUDY

B-1 Given the expanded form of a four-digit numeral, the student will write the numeral. (Expanded notation)

B-2 The student will be able to add three unlike fractions with regrouping.

B-3 The student will be able to subtract a decimal number with two decimal places from another decimal number with two decimal places with regrouping.

B-4 Given a number line showing the operation of addition of mixed numbers and an appropriate open mathematical sentence, the student will complete the mathematical sentence to describe the operation.

B-5 The student will be able to divide a three-digit whole number by a two-digit whole number with remainder.

B-6 The student will be able to multiply a proper fraction by a mixed number.

B-8 The student will be able to add quantities of time expressed as hours, minutes, and seconds with regrouping.

B-9 Given a mathematical sentence expressing the equality of the product of two symbolic quantities with the product of two other symbolic quantities, the student will be able to express true proportions using these four quantities. (Ratio and proportion)

B-10 The student will be able to subtract a three-digit whole number from a three-digit whole number with regrouping.

B-11 The student will be able to subtract mixed numbers with regrouping.

B-12 The student will be able to add two decimal numbers each with two decimal places with regrouping.

Table 1 Continued

B-13 The student will be able to multiply a three-digit whole number by a two-digit whole number.

B-14 The student will be able to divide a mixed number by a proper fraction.

B-15 The student will be able to multiply a two-digit decimal fraction by a two-digit decimal fraction.

B-16 The student will be able to divide a decimal number with one decimal place by a two-digit decimal fraction.

B-17 Given an equation in whole numbers with a missing factor, the student will be able to apply the distributive property in supplying the missing factor.

B-18 Given a sequence of whole numbers involving multiplication by a constant value, the student will be able to supply a missing number from the sequence.

B-19 Given a completed subtraction problem with missing digits in the subtrahend and the difference, the student will be able to supply the missing digits.

B-20 The student will be able to subtract quantities of liquid expressed as quarts and pints with regrouping.

C-1 The student will be able to subtract mixed numbers with regrouping.

C-2 The student will be able to multiply a decimal number with two decimal places by a decimal number with one decimal place.

C-4 The student will be able to divide a mixed number by a mixed number.

C-6 The student will be able to compute a given percent of a given whole number. (Applications)

C-7 The student will be able to add two one-digit negative integers.

C-9 Given two finite sets, the student will specify the set that is the intersection of these sets. (Intersection)

C-10 Given an open mathematical sentence involving division of positive and negative integers, the student will be able to supply the missing dividend.

C-11 Given a set of whole numbers, the student will be able to select those which are relatively prime.

Table 1 Continued

C-14 Given a set of whole numbers, the student will be able to compute their average. (Mean)

C-16 The student will be able to specify the greatest possible error of a measurement made with a given ruler.

C-18 Given the scale of a map, the student will be able to compute the distance between any two points.

C-19 Given an equation in negative integers with a missing factor, the student will be able to apply the distributive property in supplying the missing addend.

C-20 The student will be able to specify the product of two like symbolic values with different numerical exponents. (Logarithms)

TABLE 2
PMI CRITERION TESTS CLASSIFICATION AND STATISTICS

Point Biserial Range	Difficulty Range	Point Biserial With PMI Test Item	Test	Classification	No. of Items	N	KR20
.27 to .69	.80 to .94	.28	A-1	Type B	5	535	.79
.66 to .84	.50 to .63	.71	B-2 *	Type A	20	518	.97
.52 to .73	.50 to .62	.46	B-3	Type A	20	516	.94
.47 to .59	.29 to .62	.54	B-4	Type A	5	441	.77
.60 to .85	.58 to .75	.55	B-5 *	Type A	20	517	.96
.79 to .92	.38 to .46	.71	B-6 *	Type C	20	463	.98
.58 to .74	.53 to .67	.41	B-8 *	Type D	20	522	.95
.55 to .82	.06 to .17	.34	B-9	Type C	20	514	.96
.22 to .57	.78 to .85	.31	B-10	Type B	5	549	.70
.66 to .73	.37 to .46	.62	B-11	Type C	5	469	.87
.30 to .52	.53 to .65	.40	B-12	Type A	5	510	.68
.30 to .40	.52 to .67	.24	B-13	Type D	5	532	.59
.62 to .74	.41 to .52	.62	B-14	Type D	5	335	.86
.62 to .71	.39 to .47	.61	B-15	Type C	5	459	.86
.60 to .81	.29 to .40	.57	B-16 *	Type C	5	335	.90
.66 to .81	.65 to .69	.50	B-17	Type B	5	470	.88
.37 to .74	.45 to .84	.52	B-18	Type A	5	483	.81
.56 to .71	.63 to .66	.44	B-19 *	Type D	5	515	.84
.59 to .81	.24 to .50	.53	B-20	Type C	5	480	.89
.62 to .69	.54 to .64	.62	C-1 *	Type A	5	303	.85
.42 to .60	.53 to .63	.33	C-2	Type B	20	309	.89
.31 to .53	.36 to .70	.59	C-4	Unclassified	5	283	.62
.20 to .92	.05 to .55	.52	C-6	Type D	20	273	.98
.61 to .76	.70 to .78	.41	C-7	Type B	5	308	.88
.22 to .89	.19 to .75	.37	C-9	Type D	20	290	.96
.62 to .79	.42 to .51	.65	C-10	Type A	5	286	.90
.56 to .68	.14 to .30	.60	C-11	Type C	5	282	.81
.40 to .64	.45 to .78	.61	C-14	Type A	5	298	.79
.54 to .75	.33 to .43	.64	C-16	Type C	5	242	.86
.36 to .58	.54 to .86	.38	C-18 *	Type B	5	300	.72
.56 to .76	.63 to .68	.42	C-19	Type B	5	271	.87
.71 to .84	.59 to .63	.58	C-20	Type D	5	216	.92

* Distribution appears in Appendix A

TABLE 3

DESCRIPTION OF PREDICTORS
USED IN THE
PMI REGRESSION STUDY

LEVEL B

1. Operations on the number line involving addition, subtraction, multiplication, and division using whole numbers and mixed numbers. (7 items)

2. Addition of whole numbers with regrouping. (3 items)

3. Subtraction of whole numbers with regrouping. (4 items)

4. Multiplication of whole numbers including a repeated addition problem and squaring a number. (12 items)

5. Division of whole numbers with and without remainders. (9 items)

6. Addition of positive fractions and mixed numbers with and without regrouping. (6 items)

7. Subtraction of positive fractions and mixed numbers with and without regrouping. (5 items)

8. Multiplication of positive fractions and mixed numbers and items on greatest common factor, least common multiple, and lowest common denominator. (8 items)

9. Division of positive fractions and mixed numbers. (8 items)

10. Addition of decimal numbers with and without regrouping. (6 items)

11. Subtraction of decimal numbers with and without regrouping. (5 items)

12. Multiplication of decimal numbers. (8 items)

Table 3 Continued

13. Division of decimal numbers. (10 items)

14. Addition, subtraction, multiplication, and division of negative integers. (4 items)

15. Rounding numbers to the nearest 10, 100, or 1,000 and estimating answers to addition, subtraction, multiplication, and division problems. (5 items)

16. Commutative and associative properties of addition and multiplication, distributive property of multiplication over addition, and the identity elements for addition and multiplication. (10 items)

17. Determining missing elements in sequences of whole numbers or positive fractions where the sequence is determined by addition, subtraction, multiplication, or division by a constant value. (7 items)

18. Missing addends and factors involving the inverse relationships of addition, subtraction, multiplication, and division. (4 items)

19. Metric geometry involving linear measurement, area, volume, angles (90°), and line graphs. (9 items)

20. Concepts of temperature and money. (4 items)

21. Addition, subtraction, multiplication, and division of lengths, money, and liquid measures. (10 items)

22. Place value from .99999 to 999, 999, 999, expanded notation, and exponential notation. (8 items)

23. Non-base ten systems with conversions and operations and modular systems. (13 items)

24. Non-metric geometry involving recognition and classification of geometric forms. (13 items)

25. Percentage concepts involving conversion of fractions and decimal numbers to percentages and computing percentage of a given number. (6 items)

26. Set concepts including union, intersection, and Cartesion product. (5 items)

Table 3 Continued

LEVEL C

1. Trigonometry involving the definitions of sine, cosine, and tangent and an application problem. (4 items)

2. Operations on the number line involving negative integers. (4 items)

3. Addition, subtraction, multiplication, division, and squaring of whole numbers. (5 items)

4. Addition, subtraction, multiplication, and division of positive fractions and mixed numbers. (16 items)

5. Addition, subtraction, multiplication, and division of decimal numbers. (4 items)

6. Addition, subtraction, multiplication, and division of negative integers. (4 items)

7. Addition, subtraction, multiplication, and division of negative fractions and mixed numbers. (16 items)

8. Rounding numbers to the nearest 10, 100, or 1,000 and estimating answers to addition, subtraction, multiplication, and division problems. (5 items)

9. Commutative property of addition and muliplication. (5 items)

10. Associative property of addition and multiplication. (5 items)

11. Distributive property of multiplication over addition. (3 items)

12. The identity elements for addition and multiplication. (5 items)

13. Inverse relation of addition with subtraction and multiplication with division. (5 items)

14. Determining missing elements in sequences, some with alternating positive and negative numbers, where the sequence is determined by addition, subtraction, multiplication, or division by a positive or negative constant value. (4 items)

Table 3 Continued

15. Missing addends and factors involving the inverse relationships of addition, subtraction, multiplication, and division and using positive and negative whole numbers and fractions. (14 items)

16. Number theory involving prime, composite, and relatively prime numbers and odd vs. even numbers. (4 items)

17. Solution of equations. (4 items)

18. Linear measurement including preciseness, greatest possible error, relative error, and scientific notation. (5 items)

19. Metric geometry involving linear measurement, area, and volume. (6 items)

20. Non-metric geometry involving recognition and classification of geometric forms. (24 items)

21. Percentage concepts involving conversion of fractions and decimal numbers to percentages and computing percentage of a given number. (6 items)

22. Set concepts including union, intersection, Cartesian product, equivalence, subsets, empty set, disjoint sets, cardinality, and Venn diagrams. (10 items)

23. Reading values from various kinds of graphs. (7 items)

24. Statistics including finding mean, median, mode, and range. (4 items)

25. Probability including simple events, combinations, "and," and "or" problems. (4 items)

TABLE 4

SUMMARY OF RESULTS FROM THE STEPWISE REGRESSION ANALYSIS PREDICTING THE 32 PMI CRITERION TESTS FROM SUBSCORES OF THE PMI TEST
The Stepwise Analysis Was Stopped When the Next Predictor Would Contribute Less Than .003 to the Multiple Correlation

CRITERION TEST	NO. OF PREDICTORS	ORDER OF SELECTION	MULT. R	SHRUNKEN MULT.R	PMI ITEM PT.BIS.
B-1	6	4,26,5,13,3,18	.59	.58	.28
B-2	4	6,7,5,8	.86	.85	.71
B-3	8	5,11,8,10,23,1,3,15	.72	.71	.46
B-4	6	1,7,15,25,19,23	.70	.69	.54
B-5	5	5,7,4,26,3	.76	.76	.55
B-6	4	9,8,17,3	.87	.86	.71
B-8	8	6,17,5,1,2,22,15,18	.73	.72	.41
B-9	8	23,21,19,16,25,4,6,18	.26	.23	.34
B-10	7	3,5,15,16,2,20,24	.59	.58	.31
B-11	4	7,25,8,1,	.81	.81	.62
B-12	7	10,8,5,21,11,19,26	.62	.61	.40
B-13	7	5,8,4,7,6,11,3	.58	.57	.24
B-14	5	9,19,15,11,24	.85	.85	.62
B-15	4	12,8,15,3,21,17	.73	.73	.61
B-16	6	13,25,24,18,9,15	.79	.79	.57
B-17	8	8,16,5,25,15,9,12,4	.65	.64	.50
B-18	6	8,17,4,7,5,25	.69	.69	.52
B-19	8	18,5,17,10,3,24,16,4	.59	.58	.44
B-20	6	13,15,25,7,23,1	.68	.67	.53
C-1	6	4,7,24,3,17,22	.75	.74	.62
C-2	7	4,5,3,2,1,8,23	.63	.61	.33
C-4	6	4,25,9,12,17,7	.73	.72	.59
C-6	7	7,4,12,10,3,21,20	.71	.70	.52
C-7	10	4,6,2,14,15,12,1, 10,11,9	.54	.51	.41
C-9	6	5,22,2,3,20,4	.60	.59	.37
C-10	6	7,13,1,6,3,21	.78	.77	.65
C-11	9	20,16,14,13,6,17, 25,1,3	.61	.59	.60
C-14	6	4,24,15,20,25,17	.77	.76	.61
C-16	7	18,7,21,24,2,3,14	.70	.69	.64
C-18	9	4,23,12,9,1,2,22,11 5	.71	.70	.38
C-19	9	11,22,13,2,1,8,16 24,12	.68	.67	.42
C-20	10	7,17,2,23,3,1,16 25,19,13	.68	.67	.58

TABLE 5
PRI OBJECTIVES MEASURED BY AT LEAST 10 ITEMS IN THE TRYOUT STUDY
AND USED IN THE PRI RELIABILITY STUDY

A-1 The student will be able to use pronouns in sentences by choosing the correct pronoun to complete a sentence, or by substituting the correct pronoun for a noun in a sentence. (10 Items)

A-2 The student will be able to match like or unlike entities by matching given words with their definitions or synonyms. (15 Items)

A-3 The student will be able to:

1. choose the cause of a given effect in a story,
2. answer questions that require the making of inferences about a given picture,
3. answer questions that require the making of inferences about a given story, and
4. answer questions that require the drawing of conclusions from a given passage. (16 Items)

A-4 The student will be able to:

1. choose the cause of a given effect in a story,
2. answer questions that require the making of inferences about a given picture,
3. answer questions that require the making of inferences about a given story, and
4. answer questions that require the drawing of conclusions from a given passage. (16 Items)

B-1 The student will be able to recognize like vowel sounds and to distinguish between unlike vowel sounds by matching oral words with printed words, printed words with printed words, or printed words with pictures. (10 Items)

B-2 The student will be able to supply the contracted form and possessive form of given verb phrases. (12 Items)

B-3 The student will be able to:

1. match contractions used in sentences with word pairs or choose the contraction for a word pair, and
2. supply the contracted form and possessive form of given verb phrases. (20 Items)

Table 5 Continued

B-4 The student will be able to:

1. match given singular and plural words with pictures or choose designated singular or plural forms of words, and
2. match contractions used in sentences with word pairs or choose the contraction for a word pair. (15 Items)

B-5 The student will be able to indicate which of certain sentences occurred in a specific story part (e.g., last part) or choose the event that happened before or after a given event. (12 Items)

B-6 The student will be able to:

1. answer questions that require the making of inferences about a given story,
2. answer questions that require the drawing of conclusions from a given passage,
3. answer questions that require the prediction of future action from a given story, and
4. choose words or sentences that describe the feelings of a character in a given story. (17 Items)

B-7 The student will be able to choose a statement that best represents a passage or story or choose the best title for a story. (11 Items)

B-8 The student will be able to:

1. answer questions about details given in a story,
2. name the story or story part in which given events occurred, and
3. identify true statements about a given story. (18 Items)

B-9 The student will be able to indicate which of certain sentences occurred in a specific story part (e.g., last part) or choose the event that happened before or after a given event. (10 Items)

B-10 The student will be able to:

1. indicate the sense to which certain sensory images appeal, indicate which of certain sentences or phrases most appeals to a particular sense, or choose the example of sensory imagery that answers a given question or completes a given sentence, and
2. identify descriptive words and phrases from those given, or choose the most appropriate descriptive word for a given person or thing. (13 Items)

Table 5 Continued

B-11 The student will be able to:

1. draw inferences from story material by choosing responses to answer questions or to complete sentences,
2. draw logical conclusions by answering questions or choosing the best conclusion, and
3. anticipate or predict probable future actions or outcomes for story material by choosing the best answer to given questions, or by completing given sentences. (17 Items)

B-12 The student will be able to identify the theme, moral (lesson), or best title for a given selection. (10 Items)

C-1 The student will be able to add the correct affix to a word so that it fits in a given sentence or phrase. (12 Items)

C-2 The student will be able to:

1. choose words or sentences that describe character traits, attitudes, or manner of speech (e.g., spoke proudly) or answer questions about character traits, and
2. choose words or sentences that describe the feelings of a character in a given story. (13 Items)

C-3 The student will be able to:

1. recognize figures of speech by choosing a sentence that has the same meaning as a figure of speech (idiom) appearing in a given sentence, and
2. locate a simile in a given passage and choose its meaning or identify the object it describes, choose a sentence containing a simile, choose a simile to define a phrase, or identify a simile. (11 Items)

C-4 The student will be able to:

1. identify the cause of a given effect appearing in a story, and
2. identify the effect of a given action occurring in a story. (12 Items)

C-5 The student will be able to:

1. answer questions about details given in a story,
2. name the story or story part in which given events occurred, and
3. identify true statements about a given story. (13 Items)

C-6 The student will be able to:

1. answer questions about details given in a story,
2. name the story or story part in which given events occurred, and
3. identify true statements about a given story. (14 Items)

Table 5 Continued

C-7 The student will be able to choose words or sentences that describe character traits, attitudes, or manner of speech (e.g., spoke proudly) or answer questions about character traits. (11 Items)

C-8 The student will be able to:

1. recognize figures of speech by choosing a sentence that has the same meaning as a figure of speech (idiom) appearing in a given sentence,
2. locate a simile in a given passage and choose its meaning or identify the object it describes, choose a sentence containing a simile, choose a simile to define a phrase, or identify a simile, and
3. choose the definition of a metaphor, complete a given sentence with a metaphor, or identify a metaphor. (18 Items)

C-9 The student will be able to:

1. answer questions about details given in a story,
2. name the story or story part in which given events occurred, and
3. identify true statements about a given story. (10 Items)

C-10 The student will be able to identify the theme, moral (lesson), or best title for a given selection. (11 Items)

D-1 The student will be able to identify words that are compounds or choose a word to make a compound word. (13 Items)

D-2 The student will be able to:

1. choose the correct definition for given affixed words,
2. choose the correct definition of the affix in a given affixed word, and
3. choose the correct affix for an unfinished word that will complete a sentence. (24 Items)

D-3 The student will be able to:

1. recognize figures of speech by choosing a sentence that has the same meaning as a figure of speech (idiom) appearing in a given sentence,
2. locate a simile in a given passage and choose its meaning or identify the object it describes, choose a sentence containing a simile, choose a simile to define a phrase, or identify a simile, and
3. choose the definition of a metaphor, complete a given sentence with a metaphor, or identify a metaphor. (15 Items)

Table 5 Continued

D-4 The student will be able to:

1. answer questions about details given in a story,
2. name the story or story part in which given events occurred, and
3. identify true statements about a given story. (11 Items)

D-5 The student will be able to:

1. draw inferences from story material by choosing responses to answer questions or to complete sentences,
2. draw logical conclusions by answering questions or choosing the best conclusion, and
3. anticipate or predict probable future actions or outcomes for story material by choosing the best answer to given questions, or by completing given sentences. (15 Items)

D-6 The student will be able to:

1. recognize figures of speech by choosing a sentence that has the same meaning as a figure of speech (idiom) appearing in a given sentence,
2. locate a simile in a given passage and choose its meaning or identify the object it describes, choose a sentence containing a simile, choose a simile to define a phrase, or identify a simile, and
3. choose the definition of a metaphor, complete a given sentence with a metaphor, or identify a metaphor. (12 Items)

D-7 The student will be able to:

1. draw inferences from story material by choosing responses to answer questions or to complete sentences,
2. draw logical conclusions by answering questions or choosing the best conclusion, and
3. anticipate or predict probable future actions or outcomes for story material by choosing the best answer to given questions, or by completing given sentences. (11 Items)

TABLE 6

SUMMARY OF THE PRI CRITERION TEST STATISTICS AND RELIABILITY STUDY RESULTS

CRIT. TEST	GRADE	N	KR20	PT. BIS. RANGE	DIFF. RANGE	NO. ITEMS	CORRELATIONS 2-ITEMS	3-ITEMS	4-ITEMS	5-ITEMS	HIGHEST REACHED/ NO. ITEMS
A-1	1	246	.69	.09 - .56	.37 - .65	10	.31	.39	.48	.47	.48/4
	2	514	.76	.28 - .55	.84 - .97	10	.55	.65	.65	.63	.65/3
A-2	1	246	.65	-.02 - .56	.30 - .69	15	.19	.40	.44	.48	.56/7
	2	506	.79	.31 - .65	.80 - .94	15	.53	.60	.64	.70	.78/6
A-3	1	298	.77	.26 - .47	.26 - .93	16	.37	.50	.53	.53	.65/8
	2	371	.76	.23 - .53	.66 - .98	16	.43	.55	.56	.59	.64/8
A-4	1	263	.53	-.04 - .33	.32 - .80	16	.24	.25	.28	.28	.29/8
	2	408	.65	-.06 - .41	.23 - .94	16	.38	.33	.35	.37	.39/8
B-1	2	478	.80	.37 - .58	.23 - .69	10	.56	.66	.64	.60	.66/3
	3	491	.79	.26 - .57	.28 - .79	10	.53	.65	.62	.62	.65/3
B-2	2	414	.69	.18 - .48	.40 - .68	12	.28	.40	.46	.46	.56/6
	3	457	.80	.30 - .62	.56 - .84	12	.48	.57	.62	.63	.71/6
B-3	2	415	.73	.18 - .50	.40 - .84	20	.39	.52	.53	.55	.57/7
	3	460	.84	.15 - .62	.55 - .92	20	.41	.51	.53	.61	.67/6
B-4	2	373	.80	.21 - .58	.38 - .91	15	.37	.53	.61	.64	.66/6
	3	434	.77	.16 - .52	.44 - .94	15	.34	.46	.51	.57	.65/8
B-5 *	2	373	.62	.00 - .42	.29 - .89	12	.38	.46	.41	.42	.46/3
	3	434	.54	.08 - .35	.28 - .94	12	.28	.36	.38	.33	.38/4
B-6	2	373	.64	.09 - .51	.36 - .92	17	.28	.43	.51	.47	.52/8
	3	434	.64	.11 - .51	.42 - .95	17	.34	.40	.49	.48	.52/8
B-7	2	373	.60	.15 - .40	.40 - .86	11	.29	.29	.36	.38	.41/6
	3	434	.54	.07 - .39	.55 - .91	11	.30	.25	.28	.23	.28/4
B-8	2	400	.84	.11 - .62	.22 - .80	18	.54	.58	.65	.71	.74/7
	3	432	.83	.10 - .59	.36 - .91	18	.62	.64	.68	.72	.73/6
B-9	2	399	.56	-.05 - .39	.19 - .71	10	.26	.36	.39	.38	.39/4
	3	432	.59	-.03 - .40	.25 - .80	10	.27	.36	.42	.40	.42/4

* Distribution appears in Appendix A.

TABLE 6 Continued

CRIT.TEST	GRADE	N	KR20	PT. BIS. RANGE	DIFF.RANGE	NO. ITEMS	CORRELATIONS 2-ITEMS	3-ITEMS	4-ITEMS	5-ITEMS	HIGHEST REACHED/ NO. ITEMS
B-10	2	399	.65	-.08 - .55	.20 - .67	13	.38	.50	.53	.54	.54/5
	3	432	.71	.07 - .47	.36 - .81	13	.48	.57	.59	.63	.63/5
B-11	2	398	.83	.22 - .56	.26 - .84	17	.52	.53	.56	.59	.69/8
	3	432	.82	.26 - .58	.43 - .91	17	.52	.55	.55	.61	.72/8
B-12	2	389	.52	-.04 - .44	.13 - .80	10	.32	.30	.34	.28	.34/4
	3	432	.69	-.03 - .51	.13 - .87	10	.42	.48	.50	.48	.50/4
C-1	3	381	.77	.11 - .57	.17 - .70	12	.60	.63	.62	.62	.63/6
	4	511	.81	.35 - .58	.26 - .81	12	.57	.65	.67	.66	.67/4
	5	429	.82	.36 - .55	.45 - .86	12	.56	.66	.71	.69	.71/4
C-2	3	380	.77	-.02 - .58	.20 - .86	13	.59	.59	.62	.66	.66/5
	4	562	.79	.13 - .66	.31 - .91	13	.60	.65	.66	.70	.70/5
	5	401	.76	.21 - .58	.42 - .95	13	.51	.57	.58	.65	.66/6
C-3	3	379	.60	.02 - .40	.11 - .91	11	.24	.32	.36	.40	.42/6
	4	562	.64	.07 - .40	.14 - .96	11	.41	.46	.50	.49	.50/4
	5	401	.62	.04 - .42	.17 - .97	11	.41	.40	.41	.47	.47/5
C-4	3	378	.65	-.02 - .50	.13 - .87	12	.39	.32	.47	.46	.47/4
	4	562	.64	-.04 - .46	.12 - .92	12	.36	.30	.45	.46	.46/5
	5	401	.68	-.02 - .49	.14 - .95	12	.35	.34	.47	.49	.50/6
C-5	3	380	.77	.29 - .51	.36 - .82	13	.49	.56	.60	.64	.66/6
	4	562	.78	.27 - .56	.47 - .95	13	.56	.64	.64	.64	.68/6
	5	401	.79	.16 - .58	.52 - .97	13	.57	.68	.69	.71	.76/6
C-6	3	357	.79	.24 - .55	.20 - .80	14	.50	.65	.69	.70	.70/5
	4	540	.79	.20 - .60	.36 - .87	14	.52	.61	.68	.67	.68/4
	5	410	.77	.20 - .60	.40 - .93	14	.55	.57	.60	.58	.60/4
C-7	3	357	.66	.11 - .48	.12 - .76	11	.33	.41	.37	.34	.41/3
	4	540	.64	.22 - .45	.25 - .84	11	.30	.42	.36	.35	.42/3
	5	410	.68	.24 - .44	.27 - .89	11	.37	.48	.41	.37	.48/3
C-8	3	356	.69	.03 - .45	.14 - .71	18	.34	.45	.53	.52	.55/7
	4	539	.71	.01 - .48	.16 - .79	18	.37	.50	.55	.55	.60/8
	5	410	.76	.14 - .53	.20 - .77	18	.45	.57	.60	.59	.65/7

TABLE 6 Continued

CRIT. TEST	GRADE	N	KR20	PT. BIS. RANGE	DIFF. RANGE	NO. ITEMS	CORRELATIONS 2-ITEMS	3-ITEMS	4-ITEMS	5-ITEMS	HIGHEST REACHED/ NO. ITEMS
C-9	3	357	.74	.25 - .52	.34 - .79	10	.50	.59	.58	.58	.59/3
	4	539	.71	.17 - .55	.40 - .86	10	.49	.61	.60	.60	.61/3
	5	410	.71	.22 - .52	.38 - .90	10	.51	.58	.60	.61	.61/5
C-10	3	357	.66	.17 - .45	.22 - .76	11	.31	.41	.43	.51	.51/5
	4	539	.68	.16 - .47	.27 - .85	11	.35	.40	.44	.50	.51/6
	5	410	.67	.17 - .48	.32 - .85	11	.39	.43	.48	.54	.55/6
D-1 *	5	469	.84	.23 - .64	.45 - .88	13	.70	.69	.74	.77	.77/5
	6	542	.86	.26 - .67	.41 - .86	13	.70	.75	.80	-	.80/4
D-2	5	470	.64	.05 - .43	.15 - .92	24	.20	.35	.49	.55	.68/11
	6	542	.72	.01 - .50	.14 - .93	24	.31	.43	.53	.62	.71/11
D-3	5	445	.74	-.01 - .55	.14 - .88	15	.59	.59	.66	.65	.67/7
	6	508	.73	.09 - .60	.15 - .93	15	.54	.54	.59	.58	.63/7
D-4	5	445	.53	-.04 - .35	.14 - .70	11	.41	.36	.41	.41	.41/4
	6	508	.57	.03 - .39	.18 - .83	11	.37	.37	.39	.39	.39/4
D-5	5	445	.81	.24 - .60	.38 - .93	15	.37	.41	.57	.58	.67/7
	6	508	.79	.22 - .58	.50 - .94	15	.33	.39	.52	.55	.66/8
D-6	5	432	.62	.13 - .34	.14 - .83	12	.33	.39	.39	.47	.48/6
	6	511	.66	.10 - .44	.20 - .83	12	.35	.37	.39	.44	.48/6
D-7	5	432	.50	.06 - .40	.26 - .89	11	.29	.32	.31	.29	.34/6
	6	511	.56	.06 - .37	.30 - .90	11	.36	.40	.39	.36	.40/3

* Distribution appears in Appendix A.

TABLE 7

DESCRIPTION OF PREDICTORS
USED IN THE PRI REGRESSION STUDY BY
APPLICABLE CRITERION TESTS.

CRITERION TESTS A-1 and A-2

1. Choosing a correct pronoun to complete a sentence or substituting a pronoun for a noun. (4 Items)

2. Choosing the correct form of a verb to agree with the subject of a sentence. (4 Items)

3. Use of question mark, exclamation point, and quotation marks.(6 Items)

4. Combining subjects with predicates to make meaningful sentences. (4 Items)

5. Match words with their definitions or synonyms. (6 Items)

CRITERION TEST A-3

1. Recall story details and sequence of events and identify the setting of a story. (6 Items)

2. Identify character traits, feelings, and motives; identify cause and effect relationships, and draw inferences and conclusion from written or spoken material. (9 Items)

CRITERION TEST A-4

1. Identify the main idea, setting, sequence of events, and fantasy in a story. (6 Items)

CRITERION TEST B-1

1. Match vowels and rhyming word parts and identify R-controlled vowels. (12 Items)

2. Identify consonant sounds, make new words by substituting consonants, and identify the sounds of Y. (10 Items)

3. Determine the number of syllables in a word, recognize silent letters in words, and identify variant sounds of phonetic parts of words. (11 Items)

Table 7 Continued

4. Choose a correct pronoun to complete a sentence or substitute a pronoun for a noun. (4 Items)

5. Identify spelling changes caused by adding various endings to words. (4 Items)

CRITERION TEST B-2 AND B-3 (COMBINED)

1. Identify affixes and root words. (6 Items)

2. Identify contractions with the contracted words. (7 Items)

3. Distinguish between positive, comparative, and superlative forms of adjectives and recognize past, present, and future tense. (7 Items)

4. Use context to determine the meaning of words and phrases. (9 Items)

CRITERION TESTS B-4, B-5, B-6 AND B-7

1. Match words with their definitions or synonyms. (8 Items)

2. Identify homonyms and heteronyms. (6 Items)

3. Recall story details and sequence of events. (8 Items)

4. Recognize cause and effect relationships, draw inferences and conclusions from a story, and infer future action. (5 Items)

CRITERION TESTS B-8, B-9, B-10, B-11, AND B-12

1. Recall story details and sequence of events. (6 Items)

2. Identify character traits, feelings, and motives. (7 Items)

3. Identify sensory imagery and figurative language in a story. (7 Items)

4. Recognize cause and effect relationships, draw inferences and conclusions from stories, and infer future action. (9 Items)

5. Identify the main idea, fantasy, and time of occurrence of a story. (7 Items)

Table 7 Continued

CRITERION TEST C-1

1. Identify silent letters, digraphs and diphthongs, and variant phonetic parts in words and blend sounds to make new words. (12 Items)

2. Identify affixes and root words and choose definitions for affixes and affixed words. (11 Items)

3. Use of commas and exclamation marks. (6 Items)

4. Identify the referent of a pronoun, identify the meaning of given phrases in sentences, and complete sentences with appropriate phrases. (12 Items)

CRITERION TESTS C-2 AND C-4

1. Recall story details and recognize story setting. (6 Items)

2. Recognize cause and effect relationships. (4 Items)

3. Distinguish between real and make believe and recognize fables. (4 Items)

CRITERION TESTS C-6, C-7, C-8, C-9, AND C-10

1. Recall story details and sequence of events, and recognize the setting of a story. (9 Items)

2. Identify character traits, feelings, and motives and recognize sensory imagery in a story. (4 Items)

3. Recognize figurative language, simile, and metaphor in a story. (8 Items)

4. Draw conclusions from a story and infer future action. (5 Items)

5. Recognize cause and effect relationships. (4 Items)

6. Recognize the main idea and mood of a story. (6 Items)

7. Distinguish between real and make believe in a story and recognize fables. (5 Items)

TABLE 8

SUMMARY OF THE RESULTS OF THE PRI REGRESSION ANALYSIS
The Stepwise Analysis Was Stopped When the Next Predictor
Would Increase the Multiple Correlation By Less Than .003

CRIT. TEST	NO. OF PRED.	ORDER OF SEL.	MULT. R	SHRUNKEN MULT. R	4 ITEM CORRELATION	
A-1	3	1, 5, 2	.82	.82	.48	.65
A-2	3	5, 3, 4	.81	.81	.44	.64
A-3 (Cause)	2	2, 1	.69	.69	.53	.56
A-3 (Inf.)	2	2, 1	.58	.58	.53	.56
A-4	1	1 (1 pred. only)	.60	-	.28	.35
B-1	2	1, 2	.76	.76	.64	.62
B-2, 3	4	2, 1, 3, 4	.79	.79	.46 .62	.53 .53
B-4	3	3, 4, 1	.74	.74	.61	.51
B-5	3	3, 1, 4	.59	.59	.41	.38
B-6	4	3, 1, 2, 4	.55	.54	.51	.49
B-7	3	3, 4, 1	.67	.67	.36	.28
B-8	4	4, 2, 3, 1	.83	.83	.65	.68
B-9	5	4, 2, 3, 5, 1	.72	.71	.39	.42
B-10	5	1, 5, 3, 4, 2	.76	.76	.53	.59
B-11	5	4, 2, 5, 3, 1	.70	.69	.56	.55
B-12	5	4, 5, 2, 3, 1	.74	.74	.34	.50
C-1	3	2, 1, 4	.78	.78	.62 .67	.71
C-2	3	1, 3, 2	.76	.75	.62 .66	.58
C-4	3	1, 2, 3	.68	.68	.47 .45	.47
C-6	6	6, 7, 4, 1, 2, 5	.74	.74	.69 .68	.60
C-7	5	5, 7, 1, 2, 6	.68	.68	.37 .36	.41
C-8	3	3, 6, 2	.51	.51	.53 .55	.60
C-9	5	1, 7, 4, 5, 6	.72	.72	.58 .60	.60
C-10	5	5, 7, 1, 6, 4	.70	.70	.43 .44	.48